

Digital Content Session

(I did not capture the content of the presenters' slides, just tried to capture what was discussed around the slides' bullet items.)

Ben Todd – B&H publishing religious

Digital publishing

Rights to publish in digital formats

What rights do you have to digitize content? Books written a long time ago

Who should lead the charge? How do they work together?

Backlist

How to convert?

May have to buy old copy from used book store bc don't have a copy in warehouse

What format?

Xml

What to convert? How to prioritize – maybe don't sell on paper but may

Front list

Fran's comments – must first pose all the questions

Steve Kotrch – S&S

Strategy vs tactics at S&S

Digital content management and distribution

Don't own a lot of their content

Use standards

Storage standards – ink on paper bc 400-1940ad,

1940-2000 - film flats – made reprints much easier, allowed for subsidiary rights sales cause could ship the flats around; standards based on roll sizes & signatures

2000 to now - digital – but what digital?

First Iraq war – application files

2nd Iraq war – couldn't open the files from the 1st war

lessons learned – use open standards like xml & pdf, tiff, jpeg

S&S – what they archive:

Application files – quark, some can't be opened now

PDF

XML – opens up a lot of new possibilities, use it for contracts, too.

Stored in Chuck Wallace system – holds imprimatur – digital asset management system stored by ISBN, product xml, pdfs graphics, etc.

IP Intellectual properties & production – title based, brings all ISBNs together. Product info, rights info, marketing

“ISBN prime” contains the content of the book, plus other dynamic data about formats, etc.

brought into syndication server – data transformers into syndication server (work xml and rights xml) and graphics repository – To feed across firewall to service google & amazon & simon says.com– binary files, pdfs, etc.

Michael Jensen

Your digital content management solution should be defined by the scale of your company.

Strategize what you hope to do with content.

Small pub – straight forward, self built content management

Bigger pub – it is more likely you need outside tools.

One mistake they made - ISBN prime – binding specific, so should have a unique id for each title, firebrand has one. They had to retrofit because ISBN didn't work any more

Also have a file structure with directory for the number that corresponds to title with subdirectories for pdf, xml version, application files, podcasts, video, etc

Then could easily run scripts to harvest this months xml files to send to aggregators, made working with aggregation much simpler.

Reaching limits of self building at this point

Exploring firebrand as an option to make bunch of things simpler.

Production & production flow – still & likely to stay with existing in design approach. Do a lot of composition in house. Don't disrupt in house staff by putting in xml workflow. They have some maps, images, and tables, but mostly a lot of text. Xml is not best for text. Use existing flow then pdf is sent to India to turn into xml. Makes sense for their size. 2 week lag while in India. Okay with that based on lack of cost to do so.

Use text and coding.

TEI xml– industry standard added one field with bibliographic reference to save a bunch of money to develop own xml flavor.

Chose TEI initially standard & consistently structured file, can find and replace easily.

Can turn TEI xml to xhtml, etc. because of consistency. Much less \$\$\$. Good ROI.

Most decisions driven by consistency. They never name files by title of book. Use unique identifier instead. Don't have to spend tons of money on outside content service software. Not always built for publishers, for companies with lots of small files, not publishers with fewer huge files.

Store in a clean tree structure now so they can bring into a system later if they get one.

Tizra

David Durand

Tizra came out of custom sites for publishers, xml

Impact – built one piece at a time

Backlist may not be worth converting

Went to pdf

Reading doesn't change that much, or quickly.

Information seeking does, and has – starts online, must have exemplar online, must be on Google.

Purchasing habits are changing, and will continue to – people buy online now

You are in the data business

Not just books or data, but both

Ebooks are not there yet

Conservation is not enough

Get it out there now, more, better, faster.

Four C's

Cost – first thing that lead them to pdf, reviewed xml conversion & very expensive

Culture – readers, production people – let them continue with their culture

Commerce – how to sell products; pdfs – electronic paper pages

Critical mass

Questions:

For Steve – where do you store info about aspects of intellectual property
provenance – stored in IP manager that Tom Masciovecchio – designed for contract
lawyers to characterize the content to be used by the systems.

Each contract is broken down to its attributes

Served 2 purposes – needed to have versions

Go back to the contract to tie these together – territories, not rates or advances.

Married to subrights systems to see what they have doled out to get net rights
position per title

Languages, individual forms. Didn't know what distribution platform would be, but
need title info, rights, & digital files. Collect all now in repository then can go from
there.

What if objects within primary content have different rights? Working on this now.
Need to store rights info for each element.

Fran – TMM – work ID is the "ISBN prime".

Built into 7.0 – workflow of transactions between TMM and digital content
management system. Suitable for search inside book, based on dates. TMM is a
dashboard for DCMs.

Xml files – how much meta data stored in headers as part of the object, not just
about the object? Approach is to store the non-static part outside of the file. Can
insert prior to doing anything with it when needed. Should be pointed to from
central object file from content b/c it changes. Xmp has a rights piece. Allows you
to embed URIs that can be dynamic so if rights expire and photographer moves, you
can follow him for rights.

How far back in workflow is standardization? Do editors need to use xml or blue pencils? Michael hasn't standardized much. Allow people to code files up for typesetting. Not what authors and editors are good at. Don't get technology in their way to avoid resistance. Do code after the fact, not before. Let editorial process continue. S&S – introduce coding in tail end of review piece. What is chapter heading, verse, etc. They have a 17 page manual written in 2000-2002 and adopted in late 2002. It is still their standard. Could do search & destroy and turn into xml reliably. Excerpt – tag & spacers. Simple – describes structure of book.